

pandaspracticedatamining

September 4, 2025

```
[218]: import pandas as pd
[219]: df = pd.read_csv("Downloads/Datasets/iris.csv")
[220]: type(df)
[220]: pandas.core.frame.DataFrame
[221]: df
[221]:
   SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
0            5.1          3.5           1.4          0.2    Iris-setosa
1            4.9          3.0           1.4          0.2    Iris-setosa
2            4.7          3.2           1.3          0.2    Iris-setosa
3            4.6          3.1           1.5          0.2    Iris-setosa
4            5.0          3.6           1.4          0.2    Iris-setosa
..             ..
145           6.7          3.0           5.2          2.3  Iris-virginica
146           6.3          2.5           5.0          1.9  Iris-virginica
147           6.5          3.0           5.2          2.0  Iris-virginica
148           6.2          3.4           5.4          2.3  Iris-virginica
149           5.9          3.0           5.1          1.8  Iris-virginica
[150 rows x 5 columns]
[222]: df.shape
[222]: (150, 5)
[223]: df.dtypes
[223]:
SepalLengthCm    float64
SepalWidthCm     float64
PetalLengthCm    float64
PetalWidthCm     float64
Species          object
dtype: object
```

```
[224]: df.columns = ['sl', 'sw', 'pl', 'pw', 'flower_type']
df
```

```
[224]:      sl     sw     pl     pw   flower_type
0    5.1    3.5    1.4    0.2  Iris-setosa
1    4.9    3.0    1.4    0.2  Iris-setosa
2    4.7    3.2    1.3    0.2  Iris-setosa
3    4.6    3.1    1.5    0.2  Iris-setosa
4    5.0    3.6    1.4    0.2  Iris-setosa
..    ...
145   6.7    3.0    5.2    2.3 Iris-virginica
146   6.3    2.5    5.0    1.9 Iris-virginica
147   6.5    3.0    5.2    2.0 Iris-virginica
148   6.2    3.4    5.4    2.3 Iris-virginica
149   5.9    3.0    5.1    1.8 Iris-virginica
```

[150 rows x 5 columns]

```
[225]: df.dtypes
```

```
[225]: sl          float64
        sw          float64
        pl          float64
        pw          float64
        flower_type    object
       dtype: object
```

```
[226]: df.describe()
```

```
[226]:      sl            sw            pl            pw
count  150.000000  150.000000  150.000000  150.000000
mean    5.843333    3.054000    3.758667    1.198667
std     0.828066    0.433594    1.764420    0.763161
min     4.300000    2.000000    1.000000    0.100000
25%     5.100000    2.800000    1.600000    0.300000
50%     5.800000    3.000000    4.350000    1.300000
75%     6.400000    3.300000    5.100000    1.800000
max     7.900000    4.400000    6.900000    2.500000
```

```
[227]: df.head()
```

```
[227]:      sl     sw     pl     pw   flower_type
0    5.1    3.5    1.4    0.2  Iris-setosa
1    4.9    3.0    1.4    0.2  Iris-setosa
2    4.7    3.2    1.3    0.2  Iris-setosa
3    4.6    3.1    1.5    0.2  Iris-setosa
4    5.0    3.6    1.4    0.2  Iris-setosa
```

```
[228]: df.head(10)
```

```
[228]:    sl   sw   pl   pw  flower_type
 0  5.1  3.5  1.4  0.2  Iris-setosa
 1  4.9  3.0  1.4  0.2  Iris-setosa
 2  4.7  3.2  1.3  0.2  Iris-setosa
 3  4.6  3.1  1.5  0.2  Iris-setosa
 4  5.0  3.6  1.4  0.2  Iris-setosa
 5  5.4  3.9  1.7  0.4  Iris-setosa
 6  4.6  3.4  1.4  0.3  Iris-setosa
 7  5.0  3.4  1.5  0.2  Iris-setosa
 8  4.4  2.9  1.4  0.2  Iris-setosa
 9  4.9  3.1  1.5  0.1  Iris-setosa
```

```
[229]: df.tail()
```

```
[229]:    sl   sw   pl   pw  flower_type
145  6.7  3.0  5.2  2.3  Iris-virginica
146  6.3  2.5  5.0  1.9  Iris-virginica
147  6.5  3.0  5.2  2.0  Iris-virginica
148  6.2  3.4  5.4  2.3  Iris-virginica
149  5.9  3.0  5.1  1.8  Iris-virginica
```

```
[230]: df.tail(11)
```

```
[230]:    sl   sw   pl   pw  flower_type
139  6.9  3.1  5.4  2.1  Iris-virginica
140  6.7  3.1  5.6  2.4  Iris-virginica
141  6.9  3.1  5.1  2.3  Iris-virginica
142  5.8  2.7  5.1  1.9  Iris-virginica
143  6.8  3.2  5.9  2.3  Iris-virginica
144  6.7  3.3  5.7  2.5  Iris-virginica
145  6.7  3.0  5.2  2.3  Iris-virginica
146  6.3  2.5  5.0  1.9  Iris-virginica
147  6.5  3.0  5.2  2.0  Iris-virginica
148  6.2  3.4  5.4  2.3  Iris-virginica
149  5.9  3.0  5.1  1.8  Iris-virginica
```

```
[231]: df.sl
```

```
[231]: 0      5.1
 1      4.9
 2      4.7
 3      4.6
 4      5.0
 ...
145     6.7
146     6.3
```

```
147    6.5  
148    6.2  
149    5.9  
Name: sl, Length: 150, dtype: float64
```

```
[232]: df['sl']
```

```
[232]: 0      5.1  
1      4.9  
2      4.7  
3      4.6  
4      5.0  
...  
145    6.7  
146    6.3  
147    6.5  
148    6.2  
149    5.9  
Name: sl, Length: 150, dtype: float64
```

```
[233]: df.isnull()
```

```
[233]:      sl      sw      pl      pw  flower_type  
0    False  False  False  False    False  
1    False  False  False  False    False  
2    False  False  False  False    False  
3    False  False  False  False    False  
4    False  False  False  False    False  
..    ...    ...    ...    ...  
145  False  False  False  False    False  
146  False  False  False  False    False  
147  False  False  False  False    False  
148  False  False  False  False    False  
149  False  False  False  False    False
```

[150 rows x 5 columns]

```
[234]: df.isnull().sum()
```

```
[234]: sl          0  
sw          0  
pl          0  
pw          0  
flower_type  0  
dtype: int64
```

```
[235]: df.iloc[1:4, 2:4]
```

```
[235]:    pl    pw
      1  1.4  0.2
      2  1.3  0.2
      3  1.5  0.2
```

```
[236]: df1 = pd.DataFrame([[1,2],[4,5],[7,8]],
                         index = ['cobra','viper','sidewinder'],
                         columns = ['max_speed','shield'])
df1
```

```
[236]:          max_speed  shield
cobra                 1        2
viper                4        5
sidewinder            7        8
```

```
[237]: df1.loc['viper']
```

```
[237]: max_speed    4
       shield     5
       Name: viper, dtype: int64
```

```
[238]: df1.loc[['viper','sidewinder']]
```

```
[238]:          max_speed  shield
viper                 4        5
sidewinder             7        8
```

```
[239]: df.head()
```

```
[239]:    sl    sw    pl    pw  flower_type
  0  5.1  3.5  1.4  0.2  Iris-setosa
  1  4.9  3.0  1.4  0.2  Iris-setosa
  2  4.7  3.2  1.3  0.2  Iris-setosa
  3  4.6  3.1  1.5  0.2  Iris-setosa
  4  5.0  3.6  1.4  0.2  Iris-setosa
```

```
[240]: a = df.drop(0)
a.head()
```

```
[240]:    sl    sw    pl    pw  flower_type
  1  4.9  3.0  1.4  0.2  Iris-setosa
  2  4.7  3.2  1.3  0.2  Iris-setosa
  3  4.6  3.1  1.5  0.2  Iris-setosa
  4  5.0  3.6  1.4  0.2  Iris-setosa
  5  5.4  3.9  1.7  0.4  Iris-setosa
```

```
[241]: df.head()
```

```
[241]:    sl   sw   pl   pw  flower_type
      0  5.1  3.5  1.4  0.2  Iris-setosa
      1  4.9  3.0  1.4  0.2  Iris-setosa
      2  4.7  3.2  1.3  0.2  Iris-setosa
      3  4.6  3.1  1.5  0.2  Iris-setosa
      4  5.0  3.6  1.4  0.2  Iris-setosa
```

```
[242]: df.drop(0, inplace = True)
df.head()
```

```
[242]:    sl   sw   pl   pw  flower_type
      1  4.9  3.0  1.4  0.2  Iris-setosa
      2  4.7  3.2  1.3  0.2  Iris-setosa
      3  4.6  3.1  1.5  0.2  Iris-setosa
      4  5.0  3.6  1.4  0.2  Iris-setosa
      5  5.4  3.9  1.7  0.4  Iris-setosa
```

```
[243]: df.drop(df.index[0], inplace = True)
df.head()
```

```
[243]:    sl   sw   pl   pw  flower_type
      2  4.7  3.2  1.3  0.2  Iris-setosa
      3  4.6  3.1  1.5  0.2  Iris-setosa
      4  5.0  3.6  1.4  0.2  Iris-setosa
      5  5.4  3.9  1.7  0.4  Iris-setosa
      6  4.6  3.4  1.4  0.3  Iris-setosa
```

```
[244]: df.drop(df.index[3], inplace = True)
df.head()
```

```
[244]:    sl   sw   pl   pw  flower_type
      2  4.7  3.2  1.3  0.2  Iris-setosa
      3  4.6  3.1  1.5  0.2  Iris-setosa
      4  5.0  3.6  1.4  0.2  Iris-setosa
      6  4.6  3.4  1.4  0.3  Iris-setosa
      7  5.0  3.4  1.5  0.2  Iris-setosa
```

```
[245]: df.drop(df.index[[3,4]], inplace = True)
df.head()
```

```
[245]:    sl   sw   pl   pw  flower_type
      2  4.7  3.2  1.3  0.2  Iris-setosa
      3  4.6  3.1  1.5  0.2  Iris-setosa
      4  5.0  3.6  1.4  0.2  Iris-setosa
      8  4.4  2.9  1.4  0.2  Iris-setosa
      9  4.9  3.1  1.5  0.1  Iris-setosa
```

```
[246]: df.drop('sl', axis = 1)
```

```
[246]:      sl    sw    pl    pw      flower_type
 2    3.2  1.3  0.2  Iris-setosa
 3    3.1  1.5  0.2  Iris-setosa
 4    3.6  1.4  0.2  Iris-setosa
 8    2.9  1.4  0.2  Iris-setosa
 9    3.1  1.5  0.1  Iris-setosa
 ..
 ..  ...
 145   3.0  5.2  2.3 Iris-virginica
 146   2.5  5.0  1.9 Iris-virginica
 147   3.0  5.2  2.0 Iris-virginica
 148   3.4  5.4  2.3 Iris-virginica
 149   3.0  5.1  1.8 Iris-virginica
```

[145 rows x 4 columns]

```
[247]: df.tail(10)
```

```
[247]:      sl    sw    pl    pw      flower_type
 140   6.7  3.1  5.6  2.4 Iris-virginica
 141   6.9  3.1  5.1  2.3 Iris-virginica
 142   5.8  2.7  5.1  1.9 Iris-virginica
 143   6.8  3.2  5.9  2.3 Iris-virginica
 144   6.7  3.3  5.7  2.5 Iris-virginica
 145   6.7  3.0  5.2  2.3 Iris-virginica
 146   6.3  2.5  5.0  1.9 Iris-virginica
 147   6.5  3.0  5.2  2.0 Iris-virginica
 148   6.2  3.4  5.4  2.3 Iris-virginica
 149   5.9  3.0  5.1  1.8 Iris-virginica
```

```
[248]: df[df.flower_type == 'Iris-virginica'].describe()
```

```
[248]:          sl            sw            pl            pw
count  50.000000  50.000000  50.000000  50.000000
mean    6.588000  2.974000  5.552000  2.026000
std     0.635880  0.322497  0.551895  0.274650
min     4.900000  2.200000  4.500000  1.400000
25%    6.225000  2.800000  5.100000  1.800000
50%    6.500000  3.000000  5.550000  2.000000
75%    6.900000  3.175000  5.875000  2.300000
max    7.900000  3.800000  6.900000  2.500000
```

```
[249]: df.loc[0] = [1,2,3,4, 'Iris-virginica']
df.tail()
```

```
[249]:      sl    sw    pl    pw      flower_type
 146   6.3  2.5  5.0  1.9 Iris-virginica
 147   6.5  3.0  5.2  2.0 Iris-virginica
 148   6.2  3.4  5.4  2.3 Iris-virginica
```

```
149  5.9  3.0  5.1  1.8  Iris-virginica  
0     1.0  2.0  3.0  4.0  Iris-virginica
```

```
[250]: df["diff_of_sl_sw"] = df['sl'] - df['sw']  
df.head()
```

```
[250]:    sl   sw   pl   pw  flower_type  diff_of_sl_sw  
 2  4.7  3.2  1.3  0.2  Iris-setosa      1.5  
 3  4.6  3.1  1.5  0.2  Iris-setosa      1.5  
 4  5.0  3.6  1.4  0.2  Iris-setosa      1.4  
 8  4.4  2.9  1.4  0.2  Iris-setosa      1.5  
 9  4.9  3.1  1.5  0.1  Iris-setosa      1.8
```

```
[251]: df.drop('diff_of_sl_sw', axis = 1, inplace = True)
```

```
[252]: df.reset_index()
```

```
[252]:    index   sl   sw   pl   pw  flower_type  
 0        2  4.7  3.2  1.3  0.2  Iris-setosa  
 1        3  4.6  3.1  1.5  0.2  Iris-setosa  
 2        4  5.0  3.6  1.4  0.2  Iris-setosa  
 3        8  4.4  2.9  1.4  0.2  Iris-setosa  
 4        9  4.9  3.1  1.5  0.1  Iris-setosa  
 ..      ... ... ... ... ...  
 141     146  6.3  2.5  5.0  1.9  Iris-virginica  
 142     147  6.5  3.0  5.2  2.0  Iris-virginica  
 143     148  6.2  3.4  5.4  2.3  Iris-virginica  
 144     149  5.9  3.0  5.1  1.8  Iris-virginica  
 145       0  1.0  2.0  3.0  4.0  Iris-virginica
```

[146 rows x 6 columns]

```
[253]: df.reset_index(drop = True)
```

```
[253]:    sl   sw   pl   pw  flower_type  
 0  4.7  3.2  1.3  0.2  Iris-setosa  
 1  4.6  3.1  1.5  0.2  Iris-setosa  
 2  5.0  3.6  1.4  0.2  Iris-setosa  
 3  4.4  2.9  1.4  0.2  Iris-setosa  
 4  4.9  3.1  1.5  0.1  Iris-setosa  
 ..  ... ... ... ... ...  
 141 6.3  2.5  5.0  1.9  Iris-virginica  
 142 6.5  3.0  5.2  2.0  Iris-virginica  
 143 6.2  3.4  5.4  2.3  Iris-virginica  
 144 5.9  3.0  5.1  1.8  Iris-virginica  
 145 1.0  2.0  3.0  4.0  Iris-virginica
```

```
[146 rows x 5 columns]
```

```
[254]: import numpy as np
df = df.copy()
df.columns = ['sl', 'sw', 'pl', 'pw', 'flower_type']
```

```
[255]: df.iloc[2:4, 1:3] = np.nan
df.head()
```

```
[255]:    sl   sw   pl   pw  flower_type
2  4.7  3.2  1.3  0.2  Iris-setosa
3  4.6  3.1  1.5  0.2  Iris-setosa
4  5.0  NaN  NaN  0.2  Iris-setosa
8  4.4  NaN  NaN  0.2  Iris-setosa
9  4.9  3.1  1.5  0.1  Iris-setosa
```

```
[256]: df.describe()
```

```
[256]:          sl           sw           pl           pw
count  146.000000  144.000000  144.000000  146.000000
mean    5.839041    3.030556    3.865278    1.250000
std     0.916242    0.439051    1.725268    0.786985
min     1.000000    2.000000    1.000000    0.100000
25%    5.100000    2.800000    1.600000    0.325000
50%    5.800000    3.000000    4.400000    1.350000
75%    6.400000    3.300000    5.100000    1.800000
max    7.900000    4.400000    6.900000    4.000000
```

```
[257]: df.dropna(inplace = True)
df.reset_index(drop = True, inplace = True)
```

```
[258]: df.head()
```

```
[258]:    sl   sw   pl   pw  flower_type
0  4.7  3.2  1.3  0.2  Iris-setosa
1  4.6  3.1  1.5  0.2  Iris-setosa
2  4.9  3.1  1.5  0.1  Iris-setosa
3  5.4  3.7  1.5  0.2  Iris-setosa
4  4.8  3.4  1.6  0.2  Iris-setosa
```

```
[259]: df.iloc[2:4, 1:3] = np.nan
df.head()
```

```
[259]:    sl   sw   pl   pw  flower_type
0  4.7  3.2  1.3  0.2  Iris-setosa
1  4.6  3.1  1.5  0.2  Iris-setosa
2  4.9  NaN  NaN  0.1  Iris-setosa
3  5.4  NaN  NaN  0.2  Iris-setosa
```

```
4  4.8  3.4  1.6  0.2  Iris-setosa
```

```
[260]: df.sw.fillna(df.sw.mean(), inplace = True)
df.pl.fillna(df.pl.mean(), inplace = True)
df.head()
```

C:\Users\ASUS\AppData\Local\Temp\ipykernel_12040\3870255316.py:1: FutureWarning:
A value is trying to be set on a copy of a DataFrame or Series through chained
assignment using an inplace method.

The behavior will change in pandas 3.0. This inplace method will never work
because the intermediate object on which we are setting values always behaves as
a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using
'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value)
instead, to perform the operation inplace on the original object.

```
df.sw.fillna(df.sw.mean(), inplace = True)
C:\Users\ASUS\AppData\Local\Temp\ipykernel_12040\3870255316.py:2: FutureWarning:
A value is trying to be set on a copy of a DataFrame or Series through chained
assignment using an inplace method.

The behavior will change in pandas 3.0. This inplace method will never work
because the intermediate object on which we are setting values always behaves as
a copy.
```

For example, when doing 'df[col].method(value, inplace=True)', try using
'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value)
instead, to perform the operation inplace on the original object.

```
df.pl.fillna(df.pl.mean(), inplace = True)
```

```
[260]:   sl      sw      pl      pw  flower_type
 0  4.7  3.200000  1.300000  0.2  Iris-setosa
 1  4.6  3.100000  1.500000  0.2  Iris-setosa
 2  4.9  3.025352  3.898592  0.1  Iris-setosa
 3  5.4  3.025352  3.898592  0.2  Iris-setosa
 4  4.8  3.400000  1.600000  0.2  Iris-setosa
```

```
[261]: df.iloc[2:4, 1:3] = np.nan
df.head()
```

```
[261]:   sl    sw    pl    pw  flower_type
 0  4.7  3.2  1.3  0.2  Iris-setosa
 1  4.6  3.1  1.5  0.2  Iris-setosa
 2  4.9  NaN  NaN  0.1  Iris-setosa
 3  5.4  NaN  NaN  0.2  Iris-setosa
```

```
4  4.8  3.4  1.6  0.2  Iris-setosa
```

```
[262]: df_setosa = df[df.flower_type == 'Iris-setosa']
df.sw.fillna(df_setosa.sw.mean(), inplace = True)
df.pl.fillna(df_setosa.pl.mean(), inplace = True)
df.head()
```

```
C:\Users\ASUS\AppData\Local\Temp\ipykernel_12040\512207421.py:2: FutureWarning:
A value is trying to be set on a copy of a DataFrame or Series through chained
assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work
because the intermediate object on which we are setting values always behaves as
a copy.
```

For example, when doing 'df[col].method(value, inplace=True)', try using
'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value)
instead, to perform the operation inplace on the original object.

```
df.sw.fillna(df_setosa.sw.mean(), inplace = True)
C:\Users\ASUS\AppData\Local\Temp\ipykernel_12040\512207421.py:3: FutureWarning:
A value is trying to be set on a copy of a DataFrame or Series through chained
assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work
because the intermediate object on which we are setting values always behaves as
a copy.
```

For example, when doing 'df[col].method(value, inplace=True)', try using
'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value)
instead, to perform the operation inplace on the original object.

```
df.pl.fillna(df_setosa.pl.mean(), inplace = True)
```

```
[262]:    sl      sw      pl     pw  flower_type
0  4.7  3.20000  1.300000  0.2  Iris-setosa
1  4.6  3.10000  1.500000  0.2  Iris-setosa
2  4.9  3.42439  1.463415  0.1  Iris-setosa
3  5.4  3.42439  1.463415  0.2  Iris-setosa
4  4.8  3.40000  1.600000  0.2  Iris-setosa
```

```
[263]: df
```

```
[263]:    sl      sw      pl     pw  flower_type
0  4.7  3.20000  1.300000  0.2  Iris-setosa
1  4.6  3.10000  1.500000  0.2  Iris-setosa
2  4.9  3.42439  1.463415  0.1  Iris-setosa
3  5.4  3.42439  1.463415  0.2  Iris-setosa
```

```
4    4.8  3.40000  1.600000  0.2      Iris-setosa
..   ...
139   6.3  2.50000  5.000000  1.9      Iris-virginica
140   6.5  3.00000  5.200000  2.0      Iris-virginica
141   6.2  3.40000  5.400000  2.3      Iris-virginica
142   5.9  3.00000  5.100000  1.8      Iris-virginica
143   1.0  2.00000  3.000000  4.0      Iris-virginica
```

[144 rows x 5 columns]

```
[264]: df["flower_type"] = df["flower_type"].map({"Iris-setosa":1,"Iris-versicolor":2,"Iris-virginica":3})
```

```
[265]: df
```

```
 [265]:   sl      sw      pl      pw  flower_type
 0    4.7  3.20000  1.300000  0.2          1
 1    4.6  3.10000  1.500000  0.2          1
 2    4.9  3.42439  1.463415  0.1          1
 3    5.4  3.42439  1.463415  0.2          1
 4    4.8  3.40000  1.600000  0.2          1
..
139   6.3  2.50000  5.000000  1.9          ...
140   6.5  3.00000  5.200000  2.0          3
141   6.2  3.40000  5.400000  2.3          3
142   5.9  3.00000  5.100000  1.8          3
143   1.0  2.00000  3.000000  4.0          3
```

[144 rows x 5 columns]

```
[266]: df.loc[df['flower_type'] == 2]
```

```
 [266]:   sl      sw      pl      pw  flower_type
43   7.0  3.2    4.7   1.4          2
44   6.4  3.2    4.5   1.5          2
45   6.9  3.1    4.9   1.5          2
46   5.5  2.3    4.0   1.3          2
47   6.5  2.8    4.6   1.5          2
48   5.7  2.8    4.5   1.3          2
49   6.3  3.3    4.7   1.6          2
50   4.9  2.4    3.3   1.0          2
51   6.6  2.9    4.6   1.3          2
52   5.2  2.7    3.9   1.4          2
53   5.0  2.0    3.5   1.0          2
54   5.9  3.0    4.2   1.5          2
55   6.0  2.2    4.0   1.0          2
56   6.1  2.9    4.7   1.4          2
```

57	5.6	2.9	3.6	1.3	2
58	6.7	3.1	4.4	1.4	2
59	5.6	3.0	4.5	1.5	2
60	5.8	2.7	4.1	1.0	2
61	6.2	2.2	4.5	1.5	2
62	5.6	2.5	3.9	1.1	2
63	5.9	3.2	4.8	1.8	2
64	6.1	2.8	4.0	1.3	2
65	6.3	2.5	4.9	1.5	2
66	6.1	2.8	4.7	1.2	2
67	6.4	2.9	4.3	1.3	2
68	6.6	3.0	4.4	1.4	2
69	6.8	2.8	4.8	1.4	2
70	6.7	3.0	5.0	1.7	2
71	6.0	2.9	4.5	1.5	2
72	5.7	2.6	3.5	1.0	2
73	5.5	2.4	3.8	1.1	2
74	5.5	2.4	3.7	1.0	2
75	5.8	2.7	3.9	1.2	2
76	6.0	2.7	5.1	1.6	2
77	5.4	3.0	4.5	1.5	2
78	6.0	3.4	4.5	1.6	2
79	6.7	3.1	4.7	1.5	2
80	6.3	2.3	4.4	1.3	2
81	5.6	3.0	4.1	1.3	2
82	5.5	2.5	4.0	1.3	2
83	5.5	2.6	4.4	1.2	2
84	6.1	3.0	4.6	1.4	2
85	5.8	2.6	4.0	1.2	2
86	5.0	2.3	3.3	1.0	2
87	5.6	2.7	4.2	1.3	2
88	5.7	3.0	4.2	1.2	2
89	5.7	2.9	4.2	1.3	2
90	6.2	2.9	4.3	1.3	2
91	5.1	2.5	3.0	1.1	2
92	5.7	2.8	4.1	1.3	2

```
[267]: df.loc[(df['sl'] > 5.0) & (df['sw'] > 2.5) & (df['flower_type'] == 2)]
```

	sl	sw	pl	pw	flower_type
43	7.0	3.2	4.7	1.4	2
44	6.4	3.2	4.5	1.5	2
45	6.9	3.1	4.9	1.5	2
47	6.5	2.8	4.6	1.5	2
48	5.7	2.8	4.5	1.3	2
49	6.3	3.3	4.7	1.6	2
51	6.6	2.9	4.6	1.3	2

52	5.2	2.7	3.9	1.4	2
54	5.9	3.0	4.2	1.5	2
56	6.1	2.9	4.7	1.4	2
57	5.6	2.9	3.6	1.3	2
58	6.7	3.1	4.4	1.4	2
59	5.6	3.0	4.5	1.5	2
60	5.8	2.7	4.1	1.0	2
63	5.9	3.2	4.8	1.8	2
64	6.1	2.8	4.0	1.3	2
66	6.1	2.8	4.7	1.2	2
67	6.4	2.9	4.3	1.3	2
68	6.6	3.0	4.4	1.4	2
69	6.8	2.8	4.8	1.4	2
70	6.7	3.0	5.0	1.7	2
71	6.0	2.9	4.5	1.5	2
72	5.7	2.6	3.5	1.0	2
75	5.8	2.7	3.9	1.2	2
76	6.0	2.7	5.1	1.6	2
77	5.4	3.0	4.5	1.5	2
78	6.0	3.4	4.5	1.6	2
79	6.7	3.1	4.7	1.5	2
81	5.6	3.0	4.1	1.3	2
83	5.5	2.6	4.4	1.2	2
84	6.1	3.0	4.6	1.4	2
85	5.8	2.6	4.0	1.2	2
87	5.6	2.7	4.2	1.3	2
88	5.7	3.0	4.2	1.2	2
89	5.7	2.9	4.2	1.3	2
90	6.2	2.9	4.3	1.3	2
92	5.7	2.8	4.1	1.3	2

```
[268]: marks = [{'Chemistry':67, 'Physics': 45, 'Mathematics': 50, 'English': 19},  
           {'Chemistry': 90, 'Physics': 92, 'Mathematics': 87, 'English': 90},  
           {'Chemistry': 66, 'Physics': 72, 'Mathematics': 81, 'English': 72},  
           {'Chemistry': 32, 'Physics': 40, 'Mathematics': 12, 'English': 68}]  
marks_df = pd.DataFrame(marks, index = ['A', 'B', 'C', 'D'])  
marks_df
```

```
[268]:   Chemistry Physics Mathematics English  
A          67      45         50     19  
B          90      92         87     90  
C          66      72         81     72  
D          32      40         12     68
```

```
[269]: f = marks_df < 33
```

```
[270]: marks_df.mask(f, 'Fail')
```

```
[270]: Chemistry Physics Mathematics English
      A       67       45       50     Fail
      B       90       92       87     90
      C       66       72       81     72
      D     Fail       40     Fail     68
```

```
[271]: df.sort_values(by="sl")
```

```
[271]:    sl   sw   pl   pw  flower_type
143  1.0  2.0  3.0  4.0          3
6    4.3  3.0  1.1  0.1          1
35   4.4  3.2  1.3  0.2          1
31   4.4  3.0  1.3  0.2          1
34   4.5  2.3  1.3  0.3          1
...
111  7.7  2.6  6.9  2.3          ...
110  7.7  3.8  6.7  2.2          3
115  7.7  2.8  6.7  2.0          3
128  7.7  3.0  6.1  2.3          3
124  7.9  3.8  6.4  2.0          3
```

[144 rows x 5 columns]

```
[272]: df.sort_values(['sl', 'sw'], ascending = (1,0))
```

```
[272]:    sl   sw   pl   pw  flower_type
143  1.0  2.0  3.0  4.0          3
6    4.3  3.0  1.1  0.1          1
35   4.4  3.2  1.3  0.2          1
31   4.4  3.0  1.3  0.2          1
34   4.5  2.3  1.3  0.3          1
...
110  7.7  3.8  6.7  2.2          ...
128  7.7  3.0  6.1  2.3          3
115  7.7  2.8  6.7  2.0          3
111  7.7  2.6  6.9  2.3          3
124  7.9  3.8  6.4  2.0          3
```

[144 rows x 5 columns]

```
[273]: new_df=df+[5,5,7,8,2]
new_df
```

```
[273]:    sl        sw        pl   pw  flower_type
0    9.7  8.20000  8.300000  8.2      3
1    9.6  8.10000  8.500000  8.2      3
2    9.9  8.42439  8.463415  8.1      3
3   10.4  8.42439  8.463415  8.2      3
```

```

4      9.8  8.40000  8.600000  8.2          3
..    ...
139   11.3  7.50000  12.000000  9.9          5
140   11.5  8.00000  12.200000  10.0         5
141   11.2  8.40000  12.400000  10.3         5
142   10.9  8.00000  12.100000  9.8          5
143    6.0  7.00000  10.000000  12.0         5

```

[144 rows x 5 columns]

```
[274]: import numpy as np
encryp = np.sin(df)
encryp
```

```

[274]:      sl        sw        pl        pw  flower_type
0   -0.999923 -0.058374  0.963558  0.198669  0.841471
1   -0.993691  0.041581  0.997495  0.198669  0.841471
2   -0.982453 -0.279043  0.994240  0.099833  0.841471
3   -0.772764 -0.279043  0.994240  0.198669  0.841471
4   -0.996165 -0.255541  0.999574  0.198669  0.841471
..
..    ...    ...    ...    ...
139   0.016814  0.598472 -0.958924  0.946300  0.141120
140   0.215120  0.141120 -0.883455  0.909297  0.141120
141   -0.083089 -0.255541 -0.772764  0.745705  0.141120
142   -0.373877  0.141120 -0.925815  0.973848  0.141120
143   0.841471  0.909297  0.141120 -0.756802  0.141120

```

[144 rows x 5 columns]

```
[275]: output = df.apply(np.sum)
output
```

```

[275]: sl           843.100000
sw            436.448780
pl            556.526829
pw           182.100000
flower_type  296.000000
dtype: float64

```

```
[1]: !where xelatex
```

C:\Users\ASUS\AppData\Local\Programs\MiKTeX\miktex\bin\x64\xelatex.exe

```
[ ]: !jupyter nbconvert --to pdf "pandaspracticedatamining.ipynb"
```

```
[ ]:
```